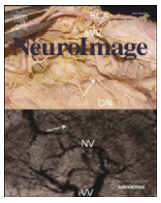
NeuroImage 118 (2015) 282–291

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg



Rapid and automatic speech-specific learning mechanism in human neocortex

Lilli Kimppa [a,*], Teija Kujala [a,b], Alina Leminen [a,c], Martti Vainio [d], Yury Shtyrov [c,e]

[a] Cognitive Brain Research Unit, Cognitive Science, Institute of Behavioural Sciences, 00014 University of Helsinki, Finland
[b] CICERO Learning, Institute of Behavioural Sciences, 00014 University of Helsinki, Finland
[c] Center of Functionally Integrative Neuroscience, Institute for Clinical Medicine, Aarhus University, 8000 Aarhus C, Denmark
[d] Phonetics and Speech Synthesis Research Group, Institute of Behavioural Sciences, 00014 University of Helsinki, Finland
[e] Centre for Cognition and Decision Making, NRU Higher School of Economics, Moscow, Russia

ARTICLE INFO

Article history:
Received 16 January 2015
Accepted 21 May 2015
Available online 11 June 2015

Keywords:
Electroencephalography
Event-related potentials
Language learning
Plasticity
Speech

ABSTRACT

A unique feature of human communication system is our ability to rapidly acquire new words and build large vocabularies. However, its neurobiological foundations remain largely unknown. In an electrophysiological study optimally designed to probe this rapid formation of new word memory circuits, we employed acoustically controlled novel word-forms incorporating native and non-native speech sounds, while manipulating the subjects' attention on the input. We found a robust index of neurolexical memory-trace formation: a rapid enhancement of the brain's activation elicited by novel words during a short (~30 min) perceptual exposure, underpinned by fronto-temporal cortical networks, and, importantly, correlated with behavioural learning outcomes. Crucially, this neural memory trace build-up took place regardless of focused attention on the input or any pre-existing or learnt semantics. Furthermore, it was found only for stimuli with native-language phonology, but not for acoustically closely matching non-native words. These findings demonstrate a specialised cortical mechanism for rapid, automatic and phonology-dependent formation of neural word memory circuits.

Introduction

Rapid acquisition of new words (sometimes called "fast mapping") is an essential ability which enables the human brain to acquire a vocabulary of thousands of words, a unique feature of human communication. Yet, the neural basis of this crucial learning mechanism is still poorly understood. In attempts to tackle this, a handful of brain imaging studies demonstrated changes in brain activation level following brief exposure to novel words. Using event-related potentials (ERPs), neural correlates of learning in adults were shown during presentation of syllable streams forming novel words by statistical extraction (Toro et al., 2005; De Diego Balaguer et al., 2007; Cunillera et al., 2009), novel words embedded in sentential contexts (Mestres-Missé et al., 2007; Borovsky et al., 2010), and in a classroom training setting (McLaughlin et al., 2004). Most importantly, a marked increase in neural activation to novel word-forms in perisylvian cortices was recently shown even after a short (~15 min) exposure to unattended previously unfamiliar spoken materials (Shtyrov et al., 2010a; Shtyrov, 2011). This evidence, interpreted as a putative correlate of rapid formation of neuronal memory circuits, has raised a number of key questions about the brain implementation of rapid word acquisition.

First, unlike the investigations conducted in active attention-demanding conditions (McLaughlin et al., 2004; Toro et al., 2005; De Diego Balaguer et al., 2007; Mestres-Missé et al., 2007; Cunillera et al., 2009; Paulesu et al., 2009; Borovsky et al., 2010; Takashima et al., 2014), data on memory-trace formation in passive exposure hinted at potential automaticity of this process[1] (Shtyrov et al., 2010a; Shtyrov, 2011; see also Saffran et al., 1997, for a behavioural study on incidental learning). This putative automaticity of learning separate novel words has not been investigated per se so far; to validate it, a more direct modulation of attention on novel spoken input is needed.

Second, it is unclear to what degree this rapid learning is speech-specific or whether it could be a general auditory learning mechanism. Although fast auditory perceptual learning effects were observed at both neural and behavioural levels after a short training, they were not found in passive exposure (Näätänen et al., 1993; Atienza et al., 2002). More specifically, no rapid learning correlates were found for non-speech signal-correlated noise stimuli in passive listening,

* Corresponding author at: Cognitive Brain Research Unit, Institute of Behavioural Sciences, PO box 9, 00014 University of Helsinki, Finland.
E-mail address: lilli.kimppa@helsinki.fi (L. Kimppa).

[1] Below, we will refer to experimental conditions where subjects are instructed to focus their attention on speech stimuli as 'attend conditions', whereas by 'passive' or 'ignore' conditions we mean those where the subjects are instructed to ignore the speech input and concentrate on an alternative primary task.

http://dx.doi.org/10.1016/j.neuroimage.2015.05.098

while acoustically similar native speech sounds exhibited fast learning-related increase in ERP amplitudes suggesting a linguistic specificity of this effect (Shtyrov, 2011). Yet, the question still remains whether such efficient learning is specific to *native* speech as it might benefit from already existing phonological action-perception circuits (Pulvermüller et al., 2012), or whether it is based on perceptual fine-tuning to *any* spoken material. Behavioural studies in children have shown less efficient fast mapping of words with non-native phonology (Rohde and Tiefenthal, 2000). However, neural underpinnings of these potentially crucial differences in learning native and non-native words have not been assessed.

Finally, and importantly, in spite of the claims that a rapid increase in brain responses to novel spoken word-forms during passive listening reflects their automatic learning and online formation of new memory circuits, no study has so far employed any behavioural learning assessment to support such claims. Retrieval of novel words after learning has been either heavily influenced by the task (e.g., list learning as in Paulesu et al., 2009), or not examined at all (Shtyrov et al., 2010a; Shtyrov, 2011). Furthermore, these latter studies included only few novel tokens (repetition of 1–4 tokens), questioning their ecological validity.

Here, we set out to comprehensively address these issues and investigate automaticity in neural memory trace build-up for words, its dependence on the presence of native speech phonology, and its linkage with behavioural learning performance, using a more natural stimulation regime with multiple spoken tokens. This was achieved by the use of an acoustically controlled set of novel word-forms (so-called "pseudo-words") with native phonology and those incorporating non-native speech sounds. Electrophysiological recordings during passive and attentive listening as well as memory recall tasks were conducted to determine the effects of attention and previous phonological exposure on rapid learning. We used the established phenomenon of an increased early brain response to familiar or recently learnt word-forms as opposed to unfamiliar ones (MacGregor et al., 2012) in order to track online memory-trace development to these different acoustically and phonologically balanced items, while manipulating attention levels on the stimulus stream.

We found a robust increase in the early (~50 ms) electrophysiological activity elicited by the same spoken syllables only when they were incorporated into native-like novel word-forms. This increase, generated by left perisylvian cortical sources, took place over a short (30 min) exposure session and did not depend on the amount of attention paid to the stimuli, suggesting automatic nature of neural memory trace build-up. Furthermore, the magnitude of this brain response increase predicted the following behavioural recall of the newly acquired items.

Materials and methods

Subjects

Twenty-two healthy right-handed (assessed according to Oldfield (1971); LQ = 82, SEM = 3.42) native speakers of Finnish (age 19–29, mean 24; 10 male) with normal hearing and no record of neurological, language or psychiatric disorders participated in the experiment. Subjects provided written informed consent and were remunerated for their participation. The study was approved by the Ethics Review Committee for Human Sciences (University of Helsinki).

Stimuli

The stimuli were created in order to precisely control and balance their acoustic and phonetic properties, while manipulating their lexical and phonological status. To this end, we used a small number of consonant-vowel (CV) syllables which were recombined to form dissyllabic (CVCV) stimuli of 3 types: (i) ten meaningful words of Finnish, (ii) ten phonotactically and phonologically legal meaningless novel word-forms (pseudo-words), and (iii) ten meaningless pseudo-words incorporating non-Finnish phonological properties. The complete set of stimuli is presented in Table 1. The native word and pseudo-word sets included identical first and second syllables, which were recombined to form lexically different items. All of the known words and native pseudo-words could only be recognised from their second syllables. This allowed us to precisely define the divergence point after which the stimulus lexicality (known word vs. meaningless novel word-form/pseudo-word) could be identified; thus the recognition points were controlled, which is important for time-locking event-related brain responses. The second syllables in non-native pseudo-words were identical to the other two sets, in order to ensure that physical features were fully matched for these critical syllables, pertinent for stimulus recognition, across all three types of stimuli. However, to give these stimuli non-native phonological properties, they were constructed by acoustically morphing together two different native syllables, resulting in novel syllables that fully retained the CV (consonant-vowel) structure and overall make-up but lacked categorisable native phonology (see Supplementary Fig. 1).

Table 1
Experimental stimuli. Stimuli comprised matched sets of meaningful Finnish words (English translations in parentheses), meaningless phonotactically and phonologically legal pseudo-words and phonologically non-native pseudo-words whose first CV syllables were produced by morphing two real existing CV-syllables (five items of each word type in a set, see the Stimuli section for details).

Known words	Pseudo-words	Non-native pseudo-words
keto (meadow)	*teto*	*pi|ta–to*
peti (bed)	*keti*	*pö|pu–ti*
poka (frame)	*puka*	*te|pa–ka*
pupu (rabbit)	*popu*	*tö|pu–pu*
teko (action)	*peko*	*pu|pä–ko*
kyky (ability)	*käky*	*te|pa–ky*
käpy (cone)	*kypy*	*pö|pu–py*
piki (pitch)	*täki*	*tö|pu–ki*
täti (aunt)	*pöti*	*pi|ta–ti*
pöpö (bug)	*pipö*	*pu|pä–pö*

All stimuli were produced from digital recordings (44.1 kHz, 16 bit) of a female native speaker of Finnish. The initial syllables were uttered in isolation and the final syllables in combination with a preceding vowel to make them unstressed, in accordance with the Finnish phonology. The syllables were selected from several utterances as matching for their fundamental frequency (F0) and vowel duration. The original syllables were normalised by the maximal peak amplitude and durations matched to 145 ms with a 5 ms fade in and a 20 ms fade out envelopes. The non-native syllables were morphed (taking into account 50% of each original native syllable) using the TANDEM–STRAIGHT algorithm (Kawahara et al., 2008). This technique decomposes sound information into source and resonator components with representation of power spectra of periodic signals, and computes an interference-free spectrum, fundamental frequency and aperiodicity. The CV structure is fully retained during the morphing procedure; yet, the phonemes do not anymore correspond to any native prototypes. This procedure offers a much more controlled way over the stimulus phonetic features than e.g. simply using a random foreign language, as it enables one to precisely dictate the make-up of the new sounds, while keeping their acoustic features and general composition on average the same as the other, native, sounds which they are based on. While some unnaturalness may result from this procedure, this ensures that the subjects have never encountered the phonemes before (which could occur when using phonemes of a foreign language). The F0 and loudness were equalised across syllables. Thereafter the F0 of the final syllables was decreased by 44 Hz and loudness by 7 dB to conform to the natural stress and prosody of Finnish speech (Suomi et al., 2003). The fundamental frequency contours of the vowels of the initial syllables were adjusted to match the contours of equivalent short Finnish vowels. The final tokens were formed by cross-splicing the initial and final syllables together with a 75 ms silent closure between the first and second

Table 2
Results of free recall and recognition tasks. Results (SEM) from ignore and attend (separated with |, respectively) conditions for native pseudo-words and known words. Mean number of correct and incorrect responses, hit rate (HR), false alarm rate (FA), and discriminability value d′ are listed.

		Correct [mean]	Incorrect [mean]	HR [%]	FA [%]	d′
Pseudo-words	Free recall	0.23 (0.09) | 1.73 (0.26)	0.95 (0.23) | 3.27 (0.54)	5 (1.83) | 35 (5.29)	NA	NA
	Recognition	1.86 (0.23) | 3.59 (0.26)	1.45 (0.32) | 1.09 (0.27)	39 (4.12) | 69 (4.43)	16 (2.83) | 13 (2.29)	0.76 (0.13) | 1.81 (0.16)
Known words	Free recall	1.50 (0.22) | 3.32 (0.23)	0.45 (0.16) | 2.05 (0.45)	30 (4.31) | 66 (4.63)	NA	NA
	Recognition	2.55 (0.25) | 4.32 (0.20)	0.91 (0.24) | 1.45 (0.23)	50 (4.86) | 80 (3.12)	11 (1.98) | 15 (2.10)	1.33 (0.19) | 2.04 (0.15)

syllables, typical of the Finnish CV onsets used. Additional target stimuli were created to be used infrequently in the attend condition by lengthening the inter-syllable gap to 145 ms. Seven native listeners rated various thereby produced tokens in order to choose the most natural sounding stimuli. Pro Tools 9 (Avid Technology, Inc., Burlington, MA, USA), Adobe Audition 3.0 (Adobe Systems Inc., San Jose, CA, USA) and Praat 5.1.45 (Boersma and Weenink, 2009) were used for stimulus production and analysis.

Procedure

The experiment was conducted in an electrically and acoustically shielded room. Stimuli were presented binaurally through headphones at 50 dB above individual hearing threshold with an average stimulus onset asynchrony (SOA) of 850 ms, jittered between 800 and 900 ms in 10 ms steps. The experiment started with an ignore condition, which was followed by an attend condition; this order was implemented in order to avoid carry-over of attention effects to the ignore session. Two subsets of five tokens from each of the three stimulus types were used in each of the conditions (such that no stimulus appeared in both conditions for any single subject). The order of the stimulus sets was counterbalanced between the subjects and conditions. The stimuli were presented in pseudo-random sequences of 15 tokens (5 of each type, equiprobably mixed) to ensure a balanced occurrence of each token type throughout the experiment. Each token was repeated 150 times in order to enable perceptual learning (Pittman, 2008; Shtyrov et al., 2010a) and to ensure a high signal-to-noise ratio of resulting ERP responses. In the ignore condition, the subjects were instructed to ignore the sounds and pay attention to a silent videofilm (cartoon without subtitles). After the ignore condition the subjects were unexpectedly asked to freely recall the stimuli by writing down what they had heard during the condition. This was followed by an equally unexpected recognition memory task, i.e. selecting the heard items from a written list of 30 alternatives (in addition to the main experimental items, this included 10 word foils and 10 pseudo-word foils which differed from the heard items by one phoneme). Conversely, in the attend condition, the subjects were instructed to pay close attention to the sounds and ignore another silent film, which was nevertheless presented to them for compatibility between conditions. Subjects' task was to press a button whenever they heard one of target stimuli which randomly occurred with 7% probability and only slightly differed from the other stimuli in having a prolonged silent closure between the two syllables. Furthermore, for the attend condition only, the subjects were instructed to memorise the words and pseudo-words heard and were informed in advance of the memory task in the end (which was otherwise identical to the unexpected test in the ignore condition). Subjects were also asked to fill in a questionnaire about the film content to assess the level of their attention in the two conditions (fifteen questions for each condition with five response alternatives including "none of the alternatives" and "I don't know"). No feedback on any of the tasks was given.

EEG recording and data pre-processing

The electroencephalogram was recorded with a 64-channel active electrode Biosemi system (Biosemi B.V., Amsterdam, Netherlands) complying with the international 10-10 system with a recording bandwidth of DC-104 Hz and 512 Hz sampling rate. The electrode site PO1 was used as a reference during the recording. Eye-movements were recorded with a horizontal and vertical electro-oculogram (EOG). Additional electrodes were placed on the mastoids and to the tip of the nose.

EEG analysis was performed using BESA Research 6.0 software (BESA Software GmbH, Münich, Germany) and Matlab R2012a (The MathWorks, Inc., Natick, MA, USA). Data were offline down-sampled to 256 Hz, bad channels were interpolated and artefact components deriving from ocular movements were removed with principal component analysis (PCA; Ille et al., 2002). Data were further filtered with a 0.5–45 Hz passband and epochs of 0–800 ms from stimulus onset were extracted. Epochs including artefacts exceeding ± 100 μV at any of the channels were removed and data were re-referenced to the mean of mastoid electrodes. The interval between the stimulus onset and the second syllable (when the stimuli could first be identified) was used as a baseline. Epochs for different tokens of each stimulus type were averaged together. Data from the first and last quarters of the experimental sessions (37 trials for each of the five tokens, resulting in 185 trials for each stimulus type in each quarter) were separated and used in further analyses. The average total number of trials after artefact rejection ranged between 150 and 164 for the "early" responses and 150–168 for the "late" ones (difference n.s.). Responses to and epochs following target stimuli and button presses were omitted from the analysis.

Source modelling

To investigate the neural underpinnings of the effect of exposure to the different types of speech input, we conducted unbiased distributed source estimations of the EEG data. BESA 6.0 software was used to run Laplacian weighted minimum norm algorithm with two iterations (LORETA; Pascual-Marqui et al., 1994). Individual subject-level source estimations were calculated for time-windows where significant ERP effects were observed (see below for details). These individual source images were averaged together for a group average image and average source maxima were used to guide the selection of voxel coordinates in the standardised Talairach space (Talairach and Tournoux, 1988). A 4-shell ellipsoidal head model was used in the source reconstruction process.

Statistical analysis

ERP data from both time periods (early and late in the exposure) were extracted for each condition and stimulus type. Response amplitudes were calculated using a 20 ms time window around peaks observed in group average waveforms; these took place at ~50 ms and ~150 ms after the stimulus divergence point (second syllable onset). These latencies correspond well to previous research showing lexical identification as early as 50 ms post stimulus uniqueness point (MacGregor et al., 2012) and later about 150 ms (Shtyrov et al., 2005). First, signals from fronto-central channels Fz and FCz, where the responses were most pronounced (Supplementary Fig. 2), were combined together. Mean amplitudes from the two response latencies were separately submitted to repeated measures analysis of variance (rmANOVA) with factors Lexicality (known word vs. pseudo-word vs. non-native pseudo-word) × Attention (ignore vs. attend) × Exposure time (early vs. late in exposure session). To investigate the effect, and its topographical

distribution in signal space more comprehensively, these were further followed up with a similar rmANOVA applied on a larger electrode array (see Supplementary data).

For statistical analysis of source data, temporal and inferior frontal source locations were chosen a priori on the basis of previous research on the importance of these areas in spoken word recognition (e.g. Pulvermüller et al., 2003; MacGregor et al., 2012; for reviews, see Hickok and Poeppel, 2007; Price, 2010), and amplitudes of neural generator activity were calculated by extracting a 21 × 21 × 21 mm voxel from each individual source estimate in each time window. Thereby collected source current densities were then submitted to rmANOVA with factors Lexicality × Source location (temporal vs. frontal source) × Hemisphere (left vs. right) × Attention × Exposure time. Whenever sphericity was violated (as assessed with Mauchly's test), multivariate ANOVA was applied as it does not assume sphericity. Bonferroni-corrected pairwise comparisons were applied post hoc for significant main effects and interactions.

Compliance with the experimental task and attention levels were assessed by comparing responses in each behavioural task and condition. Student's paired t-test was applied to the film content questionnaire. Accuracy (hit rates, HR) in free recall and recognition memory task, errors (false alarm rate, FA) in the recognition memory task, and d-prime (d′) values were calculated ($d' = Z(HR) - Z(FA)$, where Z is a Z-score transformation). The d′ is a sensitivity measure which indicates the subject's ability to discriminate the heard items from non-heard ones. In cases of perfect performance, transformations 1–1 / (2n) for perfect hits and 1 / (2n) for zero false alarms (n being the total number of hits or false alarms) were used (MacMillan and Creelman, 1991). Due to an undefinable number of possible false alarms in the free recall, the d′ was not obtained for that condition. These indicators of performance were submitted to rmANOVA with factors Attention (ignore vs. attend) × Lexicality of responses (words vs. native pseudo-words) × Validity of recollection (HR vs. FA in the recognition task and number of correct vs. incorrect response in the free recall). Hit rates were analysed using factors Memory task (free recall vs. recognition task) × Attention × Lexicality.

To determine whether increase of brain responses over time was connected to individual performance in measures of memory recall of items and attention level in the experiment, the behavioural measures and changes significant at signal and source level were taken to correlation analysis. Variables with significant 2-tailed Pearson's correlations qualifying for the assumptions of multiple linear regression were further submitted to stepwise linear regression analysis to investigate whether the degree of change in response amplitudes over the experimental session could predict the individual indexes of word memory performance and thus indicate a connection between word learning and memory retrieval. A possible effect of age was controlled in the correlation and regression analyses by entering it as a covariate. Statistical analyses were performed with SPSS Statistics 21 software (IBM Corp., Armonk, NY, United States).

Results

Exposure-related response dynamics in signal space

As soon as ~50 ms after the word divergence point, the responses to the physically identical critical syllables significantly differed between the different lexical types (see Fig. 1; Wilk's Lambda = 0.63, $F(2, 20) = 5.96$, $p = 0.009$). Post-hoc pairwise comparisons revealed that the negative response to the non-native pseudo-words was significantly smaller than that for known words and native pseudo-words (p-values < 0.008). Most importantly, there was a significant interaction ($F(2, 42) = 5.42$, $p = 0.008$) between the different lexical types and exposure time (early vs. late). This was followed by post hoc pairwise comparisons, which showed that in the beginning of the experiment (initial 25% trials) there was a significant enhancement of the responses to the known words over those to both pseudo-word types (p-values < 0.05). Additionally, legal pseudo-words with familiar phonetics and phonology elicited stronger responses than the ones with unfamiliar phonemes ($p = 0.028$). Over the course of repetitive exposure to these stimuli, however, the lexical advantage effect of word > pseudo-word disappeared and, at the end of the session, responses to native pseudo-words were even significantly larger than those to other stimuli (p-values < 0.03), and of comparable magnitude to the initial responses to known words. This phenomenon was due to native pseudo-word responses significantly increasing in amplitude during the exposure (Fig. 1; $p = 0.0255$), while those to the known words showed a decline ($p = 0.022$). The responses to the non-native pseudo-words manifested no significant amplitude changes ($p > 0.48$). Crucially, the significant increase for native pseudo-word responses did not depend on the level of attention paid to the sound stimuli (interaction Attention × Lexicality × Exposure time n.s.; $F < 1.22$, $p > 0.3$).

The second negative peak appearing at around 150 ms after the divergence point showed different effects to the earlier one (Fig. 1). A significant Attention × Lexicality interaction ($F(2, 42) = 4.78$, $p = 0.014$) suggested differential dynamics for different types of lexical input in the attend and ignore conditions. Specifically, while in the attend condition the non-native pseudo-words elicited smaller responses than the native input (p-values < 0.05), no significant differences were found in the ignore condition (p-values > 0.81). There was also a trend for a stronger negative response for attended pseudo-words than any other responses, although it did not reach full significance. In addition, a main effect of lexicality ($F(2, 42) = 8.39$, $p = 0.001$) demonstrated that, similarly to the earlier response, non-native pseudo-words elicited overall smaller responses than the other stimulus types (p-values < 0.043). No interactions of lexicality with the time of exposure were found, with an overall decrease of response amplitudes being statistically prevalent across all conditions and stimulus types ($F(1, 21) = 5.81$, $p = 0.025$). The results above were fully confirmed with a larger electrode array (see Supplementary data).

To test whether the plastic changes in these divergence-point locked responses could possibly stem from any earlier dynamics (even though it was not very likely given the acoustically and phonetically controlled stimulus design), we conducted the same analysis procedure for event-related responses time-locked to the first syllable, i.e. to the word-form onset. This showed no statistically significant effects of exposure on the responses to the initial syllables. For the first response, main effect or interactions of exposure were not significant: Lexicality × Attention × Exposure time ($F(2, 42) = 0.89$, $p = 0.418$), Lexicality × Exposure time ($F(2,42) = 1.128$, $p = 0.333$), Attention × Exposure time ($F(1,21) = 0.534$, $p = 0.473$), Exposure time ($F(1,21) = 0.741$, $p = 0.399$). Similarly, there were no significant differences between the beginning and the end of exposure for the second response peak after the first syllable onset. This confirms that the learning-related dynamics did not take place before the critical second syllable when the stimuli could be uniquely recognised.

Source analysis

As the response at 50 ms indicated significant exposure-related dynamics in signal space, these effects were followed up by individual cortical source reconstruction and analysis. As this analysis indicated (Fig. 2), source current densities in the inferior frontal and posterior temporal regions were maximal at their respective foci with atlas-based (see Source modelling) Talairach coordinates $x = \pm 38.5$, $y = 24.8$, $z = 2.8$ (inferior frontal), and $x = \pm 52.5$, $y = -31$, $z = -11.3$ (posterior temporal). Based on this, regions-of-interest (ROI) encompassing voxels within 21 mm from these maxima were defined and ROI-average source magnitudes were computed and submitted to statistical analysis.

This showed a significant interaction of Lexicality × Source location (temporal vs. frontal source) × Hemisphere (left vs. right) × Exposure

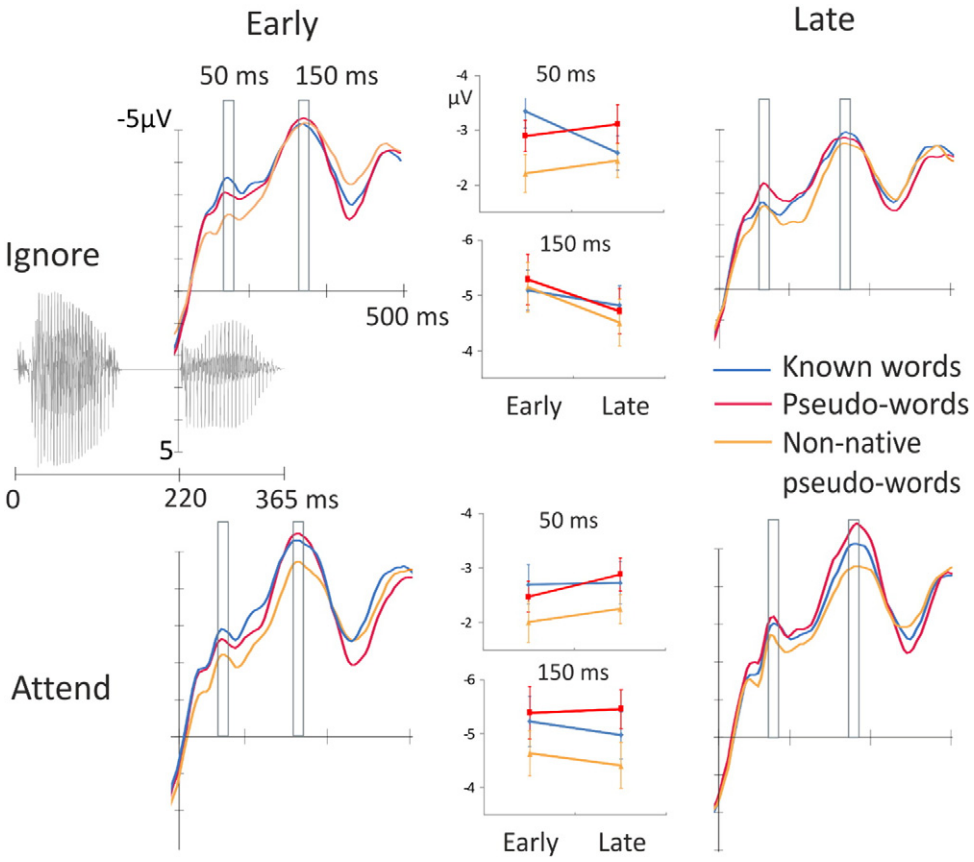


Fig. 1. Response dynamics during the course of exposure. Average ERP responses to known words (blue), pseudo-words (red) and non-native pseudo-words (orange) in the first and last quarters of the experiment (early and late) in ignore and attend conditions at the Fz–FCz ROI. Most prominent responses were found at ~50 and 150 ms latencies after the final syllable onset (divergence point). In the middle are response amplitude modulations between early and late times of exposure for each stimulus type at the two main latencies. Error bars indicate standard errors of mean (SEM). The negative-going increase of response amplitudes to pseudo-words at 50 ms was confirmed by a significant Lexicality × Exposure time interaction. The word responses manifested a significant decline in response amplitude. Responses to non-native pseudo-words showed no significant change. No interaction of attention was found. At 150 ms all lexical types attenuated over time and in the attend condition responses to native items were stronger than those to the non-native ones.

time (Wilk's Lambda = 0.736, $F(2, 20) = 3.58$, $p = 0.047$). Post hoc tests showed this to be due to the significant increment of the response to native pseudo-words for sources both in the left temporal and left inferior frontal regions (p-values < 0.05), while non-native pseudo-words demonstrated a significant decrease in the right inferior frontal source ($p = 0.026$). A frontal over temporal source activation dominance was found in the left hemisphere for words and pseudo-words early on and in the right hemisphere at the end of the session. In contrast, the dominance of the frontal source for the non-native input was found in the right hemisphere early on and in the left hemisphere in the end of the experiment (all p-values < 0.05). Furthermore, we found a significant Lexicality × Exposure time interaction ($F(2, 42) = 4.02$, $p = 0.025$) arising exclusively from an activation increase for the native pseudo-words ($p < 0.04$), while word-elicited source activation remained stable. A significant main effect for source location was found (Wilk's Lambda = 0.651, $F(1, 21) = 11.25$, $p = 0.003$) indicating that neural activity at this latency was stronger in the inferior frontal than in the temporal areas ($p = 0.003$). Neither source showed statistically significant current density differences for the suppression of the word response suggested by signal-space analysis. In addition to the inferior frontal and posterior temporal source activations, anterior temporal activations were observed (Fig. 2). These activations, however, showed no significant exposure-induced changes. The response at 150 ms, which showed attention modulation but no change due to exposure, was also followed up by source analysis; unlike the first peak, however, this did not generate statistically significant results in the inferior frontal and posterior temporal locations and thus could not be further explored.

Behavioural measures

The percentage of correct responses for the film content questionnaire demonstrated that the subjects' attention was directed as instructed in the two experimental conditions (attend mean score = 12.27, SEM = 0.25; ignore condition mean = 6.23, SEM = 0.84; $t(21) = 7.57$, $p < 0.001$). Scores from word memory tasks are shown in Table 2. Results (Fig. 3A) showed a significant interaction of Task × Lexicality for hit rate ($F(1, 21) = 8.85$, $p = 0.007$), whereby for both word types HR was higher in the recognition task than in the free recall ($p < 0.001$), and within tasks HR for known words was higher than for pseudo-words ($p < 0.018$). Furthermore, attention had a significant main effect, with better HR in the attend vs. ignore condition ($p < 0.001$). In the recognition task, comparing HR with FA yielded a significant interaction with lexicality ($F(1, 21) = 8.54$, $p = 0.008$) and attention ($F(1, 21) = 29.26$, $p < 0.001$). Hit rate was significantly higher for known words than pseudo-words ($p = 0.017$), whereas false alarm rate did not differ between word types ($p = 0.673$). Both known words and native pseudo-words had greater HR than FA ($p < 0.001$). Significantly better HR was acquired in attend condition as opposed to ignore condition ($p < 0.001$), whereas no difference between conditions was present for FA ($p > 0.73$). The d′ analysis confirmed significant main effects of attention ($F(1, 21) = 22.27$, $p < 0.001$) and lexicality ($F(1, 21) = 9.10$, $p = 0.007$) but no interaction ($F(1, 21) = 1.42$, $p > 0.24$). Discriminability was better for words than pseudo-words and in the attend than the ignore condition.

In the free recall task, there was a significant interaction of lexicality and the number of correct and incorrect responses ($F(1, 21) = 21.48$,

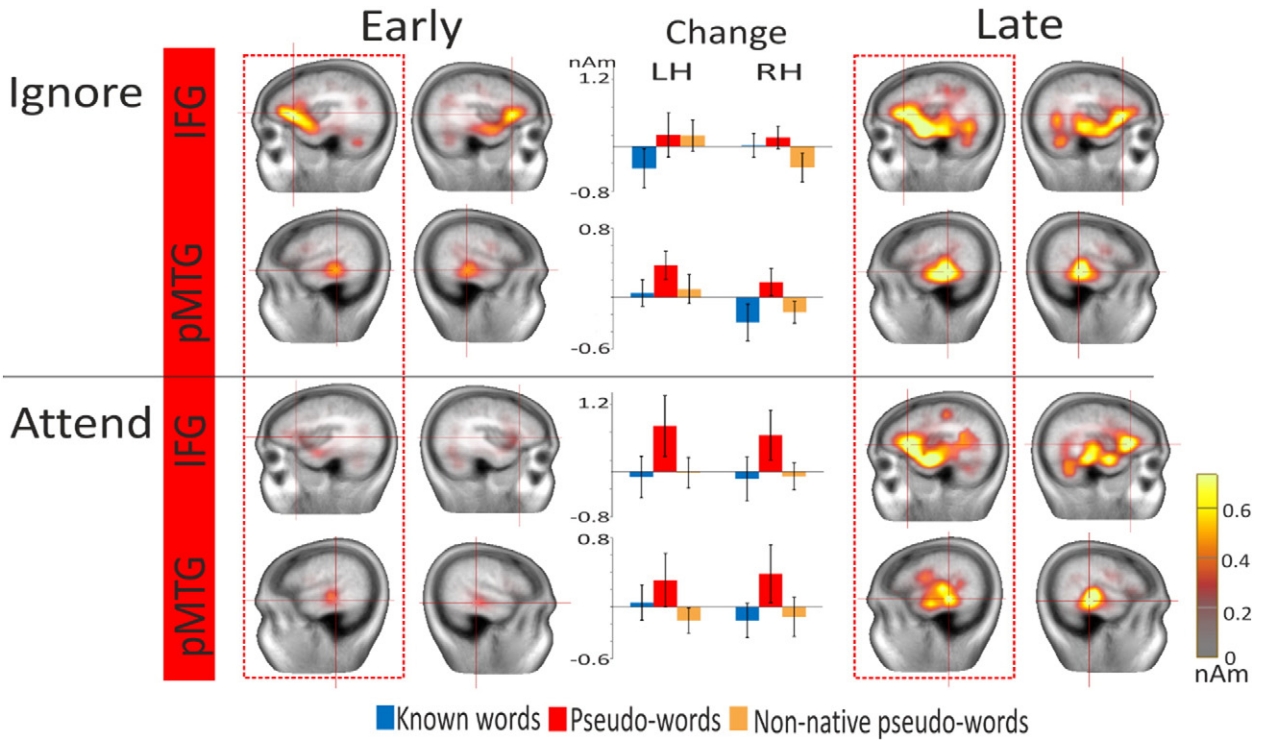


Fig. 2. Cortical source activation underlying pseudo-word learning. Average of individual LORETA source estimations of the significant response change for pseudo-words 40–60 ms post second syllable onset. Changes in source current densities between early and late exposure times for all word types in the source locations are shown in the middle. The activation increase for pseudo-words was significant in the left inferior frontal and posterior temporal locations. The crosshairs indicate the centre of the ROI used in the statistical analysis. Error bars are SEM. LH = left hemisphere, RH = right hemisphere.

$p < 0.001$) by which there were more incorrect than correct recalls of pseudo-words ($p = 0.005$) but conversely more correct than incorrect ones for known words ($p = 0.003$). At the same time the number of correctly identified known words was larger than that of pseudo-words ($p < 0.001$). The number of incorrect recalls did not significantly differ between word types ($p = 0.08$). Additionally, subjects retrieved overall more responses in the attend than ignore condition ($p < 0.001$).

In sum, these results confirm at the behavioural level the difference between ignore and attend conditions, the latter being more advantageous for richer recall. Recognising from alternatives in the recognition task produced more accurate responses than free recall. As for the recalled content, the accuracy was better for familiar as opposed to novel word forms.

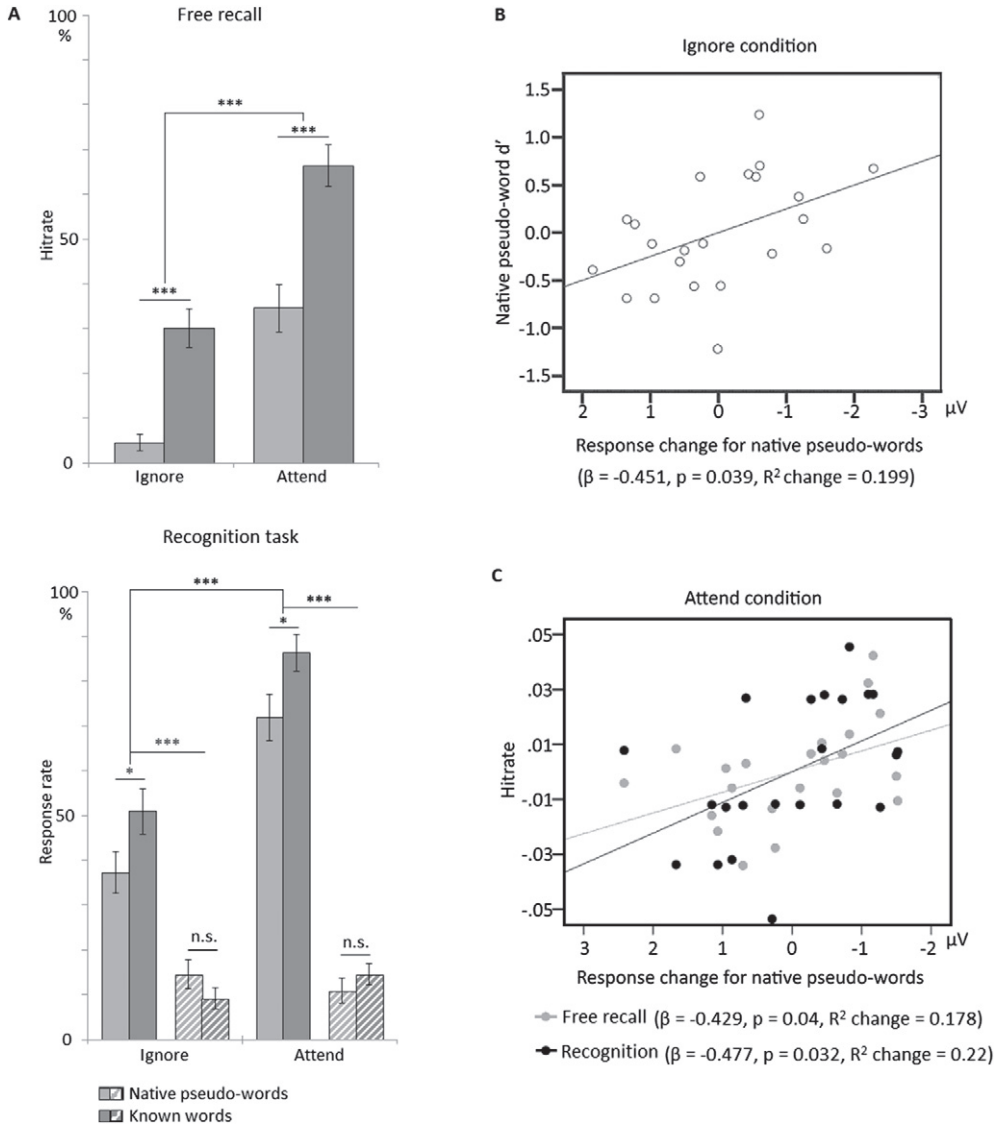
Regression analysis of behavioural and neural measures

The differences in ERPs showing a significant growth in amplitude over the exposure time were correlated with scores from memory tasks and significant correlations were followed up by linear regressions. This demonstrated clear statistical relationships between behavioural and neural measures, found for both passive and active exposure to novel lexicon. Specifically, both free overall recall ($r = -0.452$, $p = 0.04$) and recognition of pseudo-words ($r = -0.467$, $p = 0.032$) correlated with the magnitude of the pseudo-word response change at 50 ms in the attend condition. Furthermore, d′ and response change for native pseudo-words in the ignore condition also correlated significantly ($r = -0.413$, $p = 0.028$). Linear regression showed that, after removing the effect of age, the response change for native pseudo-words in the ignore condition significantly predicted the d′ of pseudo-words ($\beta = -0.451$, $p = 0.039$, R^2 change = 0.199). That is, the strength of response change for novel native items was related to their later discriminability from alternatives even when they had not been focussed on (Fig. 3B). Importantly, the response change for native pseudo-words significantly predicted the overall free recall ($\beta = -0.429$, $p = 0.04$, R^2 change = 0.178) and recognition of pseudo-words ($\beta = -0.477$, $p = 0.032$, R^2 change = 0.22) in the attend condition. Namely, a greater brain response increase for pseudo-words predicted a better ability to memorise and subsequently spontaneously recall and recognise spoken pseudo-words after attending to them (Fig. 3C).

Discussion

We determined the spatio-temporal neural dynamics linked to rapid word learning by recording the brain's electrophysiological responses to several novel word tokens with native and non-native phonology as well as to known words, while the stimuli were either ignored or attended to. A significant increase in both the event-related response amplitude and source activation in left posterior middle temporal and anterior inferior frontal areas was found in ~30 min of exposure only for the native-sounding novel word-forms. Crucially, these effects were observed as early as 50 ms after the stimulus information was sufficient for word identification. This early enhancement, reflecting neural plasticity in rapid novel word learning, did not differ between passive and attentive listening conditions. Furthermore, there was no such significant response enhancement for novel words with unfamiliar speech sounds. Unlike the novel input, responses to previously known words declined over the exposure time, most probably reflecting repetition-related suppression of neural activity (Grill-Spector et al., 2006). Moreover, using regression analysis, we found a clear relation between behavioural measures of word retrieval after exposure and the degree of neural plasticity related to learning. Below, we will briefly discuss these findings in more detail.

The learning effect found here was remarkably early and robust, the response peaking at 50 ms after the critical syllable onset (when the native stimuli diverged from their environment) across the different attention conditions. The enhanced response in this latency range has previously been linked to automatic activation of words' lexical memory traces, demonstrating that the lexical status (known word vs. pseudo-



* p < 0.05, ** p < 0.01., *** p < 0.001, n.s. = not significant

Fig. 3. Word recall performance and its relationship with response increase to pseudo-words. (A) Hit rates (solid columns) in free recall (top) and recognition task (bottom) in ignore and attend conditions. Asterisks above columns indicate significant difference between native pseudo-words (light grey) and known words (dark grey). In free recall subjects produced freely what they recalled from the heard word items followed by a recognition task where subjects selected items from a list of 30 words and pseudo-words. Memory for words was more accurate for known than pseudo-words, in the attend condition, and in the recognition task compared to the free recall. False alarm rate (striped columns) did not differ between word types or conditions. Error bars indicate SEM. (B) In the ignore condition, the change in the response to native pseudo-words served as a significant predictor for individual discriminability (d′) of pseudo-words in the recognition task. (C) In the attend condition, response change for native pseudo-words significantly predicted the overall free recall accuracy (grey dots) and accuracy specifically to pseudo-words in the recognition task (black dots). Standardised scores are presented. *p < 0.05, **p < 0.01., ***p < 0.001, n.s. = not significant.

word) of an item is neurally determined as early as 50 ms after the isolation point (MacGregor et al., 2012). The current result therefore likely reflects the online formation of new neural memory circuits within the short time of exposure to novel spoken items. Furthermore, it shows that this neural rapid learning occurs without any conscious effort and that this process becomes manifest near-instantaneously when the available information allows for word-form identification. As compared to the initial reports of passive rapid learning effects at 80–120 ms post lexical divergence point (Shtyrov et al., 2010a; Shtyrov, 2011), we here show different response dynamics: an earlier learning-related peak at ~50 ms with a later response at ~150 ms. This divergence could be due to differences in stimulus structures between studies (mono-, disyllabic) as well as the massive repetition of the same very small number (1–4) of stimuli in the previous studies, where they were presented in a rather unnatural oddball sequence at strictly fixed SOAs. These factors possibly led to suppression of the earlier response in those experiments. Here, 30 different pseudo-randomly mixed items were presented with equal probability at jittered SOA, which resulted in a somewhat more natural presentation mode with more stimulus variability that could, at least to a degree, counteract the repetition-related inhibition of the earliest lexical response phase.

An important advantage of the current study compared to previous ones that involved only one kind of tasks (cf. Toro et al., 2005; Saffran et al., 1997) is the systematic modulation of attention on the speech input: the subjects were clearly instructed to either completely ignore the spoken input and concentrate on a primary task in the visual modality, or to closely attend to speech stimuli, while the experimental conditions were otherwise kept identical. Strikingly, this controlled manipulation allowed us to reveal the neural similarity of rapid learning phenomena between situations when spoken input is ignored vs. when it is attended and memorised. This strongly suggests that neural learning of novel spoken materials is highly automatised, at least at its

initial lexicalisation stages. This automaticity could be what enables us to quickly and efficiently develop large vocabularies as children, and continue increasing their size later in life.

As a side note, another noteworthy difference between this and many of the previous studies involving attend conditions (Toro et al., 2005; McNealy et al., 2006; De Diego Balaguer et al., 2007; Cunillera et al., 2009), is that they required extraction of novel syllable combinations from a continuous speech stream by manipulating statistical probabilities of these combinations, which are sometimes even distributed over longer time periods. This statistical learning is different from the exposure to clearly distinct isolated word-forms inter-mixed with real words, which we employed here. To what extent the automaticity we report is specific to this type of word exposure, and/or whether it can also be found in other paradigms, remain to be investigated in future studies. Furthermore, a direct comparison of the effects found here and the results in e.g. De Diego Balaguer et al. (2007) and Cunillera et al. (2009), would be somewhat misleading, as previous studies typically quantified later responses, calculated from the onset of the complete word, whereas we time-locked our response to the critical identification points, in order to obtain more precise temporal information on the dynamics of these neurolinguistic processes.

As for the second response phase at 150 ms, there was a trend for a larger native pseudo-word > word response in attend condition (Fig. 1). Similar to this, previous studies found stronger responses for attended pseudo-words than known words at 120–150 ms and later (Garagnani et al., 2009; Shtyrov et al., 2010b). These were explained by resource demanding processes of finding a lexical match for attended yet unfamiliar word-like input. The most crucial finding here, however, is that the second response phase showed no learning-related increase over the course of exposure. Thus, the effect of attention appears to arise temporally later than the initial learning effect, and in itself does not significantly affect the early stage of rapid neural memory trace formation for novel words. Any later effects, e.g. in the N400 range, were not obviously present in the current study. Those are more likely to arise in a situation where integration of the novel stimulus with context is required (Mestres-Missé et al., 2007); as our stimuli were presented outside any sentential context, N400 effects were not explicitly predicted.

In contrast to the automatic response enhancement for the phonologically native novel word-forms, no such effect was found for the neural processing of pseudo-words with embedded non-native speech sounds. Crucially, while being comparable acoustically and having the same overall consonant–vowel structure, the non-native input included novel phoneme morphs, which could essentially be interpreted as unfamiliar previously unencountered phonemes with no corresponding phonetic/phonological representations in the subjects' phonetic inventory. Such absence of pre-existing phonological memory circuits therefore appears to prevent rapid learning of non-native word-forms. Moreover, the later response showed a clear suppression for these stimuli, indicating that the uncategorisable speech sounds were subject to less elaborate processing than the native ones. Further, the second phase of non-native pseudo-word responses showed an attentional modulation: no difference in amplitude from other stimulus types in the ignore condition but a smaller response magnitude in the attend condition. Attention to the novel non-native input thus counteracted the typically observed response enhancement for attended pseudo-words.

Word learning in a second language with unfamiliar phonology is thus not automatised; it is likely to be considerably slower and require higher level of attention than that for native words. This is also supported by previous observations of smaller brain response to words of a second language in learners with low proficiency in contrast to native-like activation in high proficient learners (Ojima et al., 2005; Wong et al., 2007). The critical prerequisite for fast mapping of new lexical items therefore seems to be an existing neural circuit for their low-level phonological constituents which they can be mapped onto, facilitating the assembly of those constituents to form new lexical units.

Still, caution is warranted in interpreting these results on different neural strategies for learning native and non-native speech. First, while the design ensured that the relevant neurolexical processes could only fully unravel at the beginning of the second syllable, there was, inevitably, a necessary acoustic difference between the native and non-native item sets, localised at the first syllable (in order to have acoustically identical second syllables and thus minimise acoustic confounds). Whereas we found no significant exposure-related ERP effects following the first syllable onset, we cannot fully exclude that some sort of rapid learning dynamics may arise for non-native sounds as well. This needs to be explored in future studies which can for example apply similar morphing manipulations on the critical second syllable or on both the first and the second syllables of disyllabic items, or use monosyllabic stimuli of different types. Second, there is still a theoretical possibility of assimilation of novel phonemes into existing native categories, a process possible in second language learning (Best and Tyler, 2007). This, however, cannot be validated in the current result. While the result of more incorrect responses in the free recall in attend condition could be interpreted as subjects trying to formulate or spell out the non-native items using native phonology/orthography, this is impossible to verify as our memory tasks could not accommodate any provision for an unambiguous self-report of non-native items.

Critically, our study demonstrates a clear statistical association between individual electrophysiological learning measures and behavioural outcomes, i.e. memory recall of new words, which was condition-specific. That is, the magnitude of neural response increase to attended pseudo-words significantly predicted their consecutive behavioural recall. Moreover, the subject's ability to discriminate the newly acquired items from unfamiliar fillers was predicted by the native pseudo-word ERP increase in the ignore condition. These findings argue for item specificity between the behavioural and neural learning.

Since no learning effect was found for the non-native items at the neural level, no learning can be assumed behaviourally either. This, however, as we already mentioned above, cannot be fully elucidated here, since the non-native items could not be transcribed using conventional memory tests; future studies could address this using e.g. auditory recognition tasks. For the native phonology, however, our overall results clearly link the rapid neural plastic changes to word retrieval, supporting the view of rapid learning as a basis for longer-term memory trace formation for new words (Shtyrov, 2012). The behavioural results also show better accuracy for attentive listening, indicating that, while the initial lexicalisation can be automatic at neural level, attention still has a clear influence on behavioural performance outcomes.

The enhancement in source activity for novel words originated from the left IFG, with peak activations in its anterior part (~BA45), and the posterior MTG (~BA21). Although caution is needed with respect to the precision of source localisation of EEG data, our source reconstruction could nevertheless show these two left-lateralised locations as the prominent sites for native novel word learning. Structural (Catani and Jones, 2005; Glasser and Rilling, 2008) and functional (Pulvermüller et al., 2006; Shtyrov and Pulvermüller, 2007; López-Barroso et al., 2013) connections between these areas are well established, strongly suggesting that they form the core of language networks in the brain. Indeed, their activation was also found in studies of novel word learning relying on phonology (Paulesu et al., 2009; Takashima et al., 2014) and semantics (Mestres-Missé et al., 2008, 2009).

An increase in the left posterior temporal and inferior frontal activity was previously shown to predict the retrieval of new words from long-term memory over a 10 month period (Hultén et al., 2009). Posterior MTG is proposed to represent a lexical interface (Hickok and Poeppel, 2007; Gow, 2012), activated by lexico-semantic processing (Davis and Gaskell, 2009) whereas the left anterior IFG was shown to be critical in lexical decisions (Heim et al., 2005; Xiao et al., 2005), phonological analysis of novel words and retrieval from memory (Clark and Wagner, 2003), and articulatory rehearsal (Paulesu et al., 1993).

The current model did not reveal statistically reliable sources of attention-related effects found in signal space at the later peak; future studies using methods with higher spatial resolution, such as simultaneous MEG/EEG recording combined with individual structural MRI, could explore the neuroanatomical correlates of extra-linguistic factors (e.g. attention) related to learning.

Conclusions

To conclude, this study documents a neural correlate of rapid word learning, a crucial human skill that plays a defining role in our social communication abilities. The learning effect is temporally early, emerging already at ~50 ms after the spoken input can be identified, and demonstrates a highly automatised process resilient to attention modulation. Importantly, the strength of the brain response increase predicts the subsequent recall and recognition of the newly learnt words. However, this fast automatic neural memory trace build-up is restricted to words with familiar native-language phonology, benefiting from pre-existing perception-articulation links in the brain. This, in turn, implies different neural strategies for learning new words in the native and non-native languages. The behavioural results together with the lack of attention effect in the early learning-related response phase suggest that attention affects word processing only somewhat later on, after the initial automatic stages of neurolexical activation are complete. The primary cortical circuits underlying rapid novel word learning comprise structures of the left posterior middle temporal and inferior frontal cortices, which allow harnessing both the phonological and lexico-semantic networks critical for mapping novel speech input onto neural memory representations.

The current findings shed new light on the fascinating facility of the human brain to rapidly build new memory traces for linguistic elements, enabling successful human communication. The automatic neural mechanism described here is what likely underpins the ease in learning new words of the native tongue even in adulthood. On the other hand, the specificity of this rapid learning to native phonology can explain the difficulties experienced by adults when acquiring new languages. The adult neural language system seems to be highly fine-tuned and receptive to native language phonology; future research can explore the development of this mechanism in infancy and childhood.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.05.098.

Funding

This work was supported by the Academy of Finland Grant 276414, Science Foundation of the University of Helsinki, Ella and Georg Ehrnrooth Foundation, Doctoral Programme of Psychology, Learning and Communication (University of Helsinki) and the Lundbeck Foundation (Denmark) (Grant 2013-12951).

Acknowledgments

We thank Tommi Makkonen for technical assistance.

Conflict of interest

None declared.

References

Atienza, M., Cantero, J.L., Dominguez-Marin, E., 2002. The time course of neural changes underlying auditory perceptual learning. Learn. Mem. 9, 138–150.
Best, C.T., Tyler, M.D., 2007. Nonnative and second-language speech perception: commonalities and complementarities. Language Experience in Second Language Speech Learning: In Honor of James Emil Flegepp. 13–34.
Boersma, P., Weenink, D., 2009. Praat: doing phonetics by computer [computer program] retrieved from. http://www.praat.org.
Borovsky, A., Kutas, M., Elman, J., 2010. Learning to use words: event-related potentials index single-shot contextual word learning. Cognition 116, 289–296.
Catani, M., Jones, D.K., 2005. Perisylvian language networks of the human brain. Ann. Neurol. 57, 8–16.
Clark, D., Wagner, A.D., 2003. Assembling and encoding word representations: fMRI subsequent memory effects implicate a role for phonological control. Neuropsychologia 41, 304–317.
Cunillera, T., Càmara, E., Toro, J.M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., Pujol, J., Rodríguez-Fornells, A., 2009. Time course and functional neuroanatomy of speech segmentation in adults. Neuroimage 48, 541–553.
Davis, M.H., Gaskell, M.G., 2009. A complementary systems account of word learning: neural and behavioural evidence. Philos. Trans. R. Soc. Lond. B Biol. Sci. 364, 3773–3800.
De Diego Balaguer, R., Toro, J.M., Rodriguez-Fornells, A., Bachoud-Lévi, A.C., 2007. Different neurophysiological mechanisms underlying word and rule extraction from speech. PLoS ONE 2, e1175.
Garagnani, M., Shtyrov, Y., Pulvermüller, F., 2009. Effects of attention on what is known and what is not: MEG evidence for functionally discrete memory circuits. Front. Hum. Neurosci. 3, 1–12.
Glasser, M.F., Rilling, J.K., 2008. DTI tractography of the human brain's language pathways. Cereb. Cortex 18, 2471–2482.
Gow Jr., D.W., 2012. The cortical organization of lexical knowledge: a dual lexicon model of spoken language processing. Brain Lang. 121, 273–288.
Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn. Sci. 10, 14–23.
Heim, S., Alter, K., Ischebeck, A.K., Amunts, K., Eickhoff, S.B., Mohlberg, H., Zilles, K., von Cramon, D.Y., Friederici, A.D., 2005. The role of the left Brodmann's areas 44 and 45 in reading words and pseudowords. Cogn. Brain Res. 25, 982–993.
Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402.
Hultén, A., Vihla, M., Laine, M., Salmelin, R., 2009. Accessing newly learned names and meanings in the native language. Hum. Brain Mapp. 30, 976–989.
Ille, N., Berg, P., Scherg, M., 2002. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. J. Clin. Neurophysiol. 19, 113–124.
Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T., Banno, H., 2008. TANDEM–STRAIGHT: a temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation. P IEEE Int Conf Acoust Speech Signal Proc. 2008 3933–3936.
López-Barroso, D., Catani, M., Ripollés, P., Dell'Acqua, F., Rodríguez-Fornells, A., de Diego-Balaguer, R., 2013. Word learning is mediated by the left arcuate fasciculus. Proc. Natl. Acad. Sci. 110, 13168–13173.
MacGregor, L.J., Pulvermüller, F., van Casteren, M., Shtyrov, Y., 2012. Ultra-rapid access to words in the brain. Nat. Commun. 3, 711.
Macmillan, N.A., Creelman, C.D., 1991. Detection Theory: A User's Guide. Cambridge University Press, Cambridge (UK).
McLaughlin, J., Osterhout, L., Kim, A., 2004. Neural correlates of second-language word learning: minimal instruction produces rapid change. Nat. Neurosci. 7, 703–704.
McNealy, K., Mazziotta, J.C., Dapretto, M., 2006. Cracking the language code: neural mechanisms underlying speech parsing. J. Neurosci. 26, 629–7639.
Mestres-Missé, A., Rodriguez-Fornells, A., Münte, T.F., 2007. Watching the brain during meaning acquisition. Cereb. Cortex 17, 1858–1866.
Mestres-Missé, A., Càmara, E., Rodriguez-Fornells, A., Rotte, M., Münte, T.F., 2008. Functional neuroanatomy of meaning acquisition from context. J. Cogn. Neurosci. 20, 2153–2166.
Mestres-Missé, A., Münte, T.F., Rodriguez-Fornells, A., 2009. Functional neuroanatomy of contextual acquisition of concrete and abstract words. J. Cogn. Neurosci. 21, 2154–2171.
Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., Paavilainen, P., 1993. Development of a memory trace for a complex sound in the human brain. Neuroreport 4, 503–506.
Ojima, S., Nakata, H., Kakigi, R., 2005. An ERP study of second language learning after childhood: effects of proficiency. J. Cogn. Neurosci. 17, 1212–1228.
Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.
Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49–65.
Paulesu, E., Frith, C.D., Frackowiak, R.S., 1993. The neural correlates of the verbal component of working memory. Nature 362, 342–345.
Paulesu, E., Vallar, G., Berlingeri, M., Signorini, M., Vitali, P., Burani, C., Perani, D., Fazio, F., 2009. Supercalifragilisticexpialidocious: how the brain learns words never heard before. Neuroimage 45, 1368–1377.
Pittman, A.L., 2008. Short-term word-learning rate in children with normal hearing and children with hearing loss in limited and extended high-frequency bandwidths. J. Speech Lang. Hear. Res. 51, 785–797.
Price, C.J., 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann. N. Y. Acad. Sci. 1191, 62–88.
Pulvermüller, F., Shtyrov, Y., Ilmoniemi, R., 2003. Spatiotemporal dynamics of neural language processing: an MEG study using minimum-norm current estimates. Neuroimage 20, 1020–1025.
Pulvermüller, F., Shtyrov, Y., Ilmoniemi, R.J., Marslen-Wilson, W.D., 2006. Tracking speech comprehension in space and time. Neuroimage 31, 1297–1305.
Pulvermüller, F., Kiff, J., Shtyrov, Y., 2012. Can language–action links explain language laterality?: an ERP study of perceptual and articulatory learning of novel pseudowords. Cortex 48, 871–881.
Rohde, A., Tiefenthal, C., 2000. Fast mapping in early L2 lexical acquisition. Studia Linguistica 54, 167–174.

Saffran, J.R., Newport, E.L., Aslin, R.N., Tunick, R.A., Barrueco, S., 1997. Incidental language learning: listening (and learning) out of the corner of your ear. Psychol. Sci. 8, 101–105.
Shtyrov, Y., 2011. Fast mapping of novel word forms traced neurophysiologically. Front. Psychol. 2, 340.
Shtyrov, Y., 2012. Neural bases of rapid word learning. Neuroscientist 18, 312–319.
Shtyrov, Y., Pulvermüller, F., 2007. Early MEG activation dynamics in the left temporal and inferior frontal cortex reflect semantic context integration. J. Cogn. Neurosci. 19, 1633–1642.
Shtyrov, Y., Pihko, E., Pulvermüller, F., 2005. Determinants of dominance: is language laterality explained by physical or linguistic features of speech? Neuroimage 27, 37–47.
Shtyrov, Y., Nikulin, V.V., Pulvermüller, F., 2010a. Rapid cortical plasticity underlying novel word learning. J. Neurosci. 30, 16864–16867.
Shtyrov, Y., Kujala, T., Pulvermüller, F., 2010b. Interactions between language and attention systems: early automatic lexical processing? J. Cogn. Neurosci. 22, 1465–1478.
Suomi, K., Toivanen, J., Ylitalo, R., 2003. Durational and tonal correlates of accent in Finnish. J. Phon. 31, 113–138.
Takashima, A., Bakker, I., van Hell, J.G., Janzen, G., McQueen, J.M., 2014. Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. Neuroimage 84, 265–278.
Talairach, J., Tournoux, P., 1988. Co-planar Stereotaxic Atlas of the Human Brain. 3-Dimensional Proportional System: An Approach to Cerebral Imaging. Thieme, Stuttgart (Germany).
Toro, J.M., Sinnett, S., Soto-Faraco, S., 2005. Speech segmentation by statistical learning depends on attention. Cognition 97, B25–B34.
Wong, P., Perrachione, T.K., Parrish, T.B., 2007. Neural characteristics of successful and less successful speech and word learning in adults. Hum. Brain Mapp. 28, 995–1006.
Xiao, Z., Zhang, J.X., Wang, X., Wu, R., Hu, X., Weng, X., Tan, L.H., 2005. Differential activity in left inferior frontal gyrus for pseudowords and real words: an event-related fMRI study on auditory lexical decision. Hum. Brain Mapp. 25, 212–221.